A Mobile Application for Early Diagnosis of Pneumonia in the Rural context

Unais Sait
Faculty of Architecture and Design
PES University
Bangalore, India
unaissait@pes.edu

Sanjana Shivakumar
School of Architecture
Ramaiah Institute of Technology
Bangalore, India
sanjanashivakumar17@gmail.com

Gokul Lal K V
Department of E.E.E
EPCET
Bangalore, India
gokuintheuniverse@gmail.com

Tarun Kumar
Center for Product Design and Manufacturing
Indian Institute of Science
Bangalore, India
tarunkumar@iisc.ac.in

Vandana D Ravishankar
Department of E.C.E
PES University
Bangalore, India
vandu6767@gmail.com

Kriti Bhalla
School of Architecture
Ramaiah Institute of Technology
Bangalore, India
kritibhalla@msrit.edu

***Abstract*— Pneumonia is a respiratory infection resulting in inflammation of the lungs. The causes of this infectious disease could be attributed to viruses, bacteria or fungi. One of the many ways of detecting the disease is by a chest X-ray of the patient. The rural population in developing nations have limited access to doctors, medical diagnostic facilities, and hospitals. Hence, diagnosis is delayed resulting in adverse consequences. This paper is an attempt to design and develop a smartphone-based application (app) for the preliminary detection of pneumonia using X-ray images. The app is based on machine learning which identifies pneumonia, using a chest X-ray image of a patient with a '*MobileNets*' model, trained on thousands of X-ray images of known cases of pneumonia. The app has been developed on Android Studio, incorporating TensorFlow library. The patient's chest X-ray is scanned and uploaded to the app using the smartphone camera. Additionally, an e-diagnosis facility is integrated into the app where qualified medical practitioners' advice is taken on the obtained results. A *breathing pattern recorder module* is developed, which, in future, could be integrated into the smartphone app to increase its accuracy in prediction.**

Keywords—X-ray, pneumonia, breathing pattern, TensorFlow, MobileNets, Machine Learning

I. Introduction

Pneumonia is a type of lung infection, caused by bacteria and bacteria-like organisms (*Streptococcus pneumoniae, Haemophilus influenza, Staphylococcus aureus, Mycoplasma pneumoniae*), virus (*influenza type A and type B, respiratory syncytial virus* (RSV)), fungi (*coccidioides fungus*) in surrounding environments [1],[2]. Exposure and inhalation of the contaminated air in these environments ultimately leads to inflammation, and fluid filling in the lungs, in turn, reducing oxygen flow to the bloodstream [1],[3]. Aspiration pneumonia, pneumonia acquired by patients in hospitals (via contact with ventilators, instruments) are other categories of pneumonia. Viruses are the primary cause of pneumonia in children under five years [2]. Children, infants, elderly, people with weakened immune systems, and people with severe alcohol misuse have an increased risk.

It remains the leading infection-based cause of death among children below five years of age, approximately killing 100 children per hour [4]. Out of 5.6 million under-five deaths, pneumonia has accounted for 16 per cent, killing 880,000 children in 2016 [9]. In India, the Infant mortality rate is 39 deaths/1,000 live births (fig.1), i.e., 38 deaths/1,000 live births among males and 40 deaths/1,000 live births among females [5].

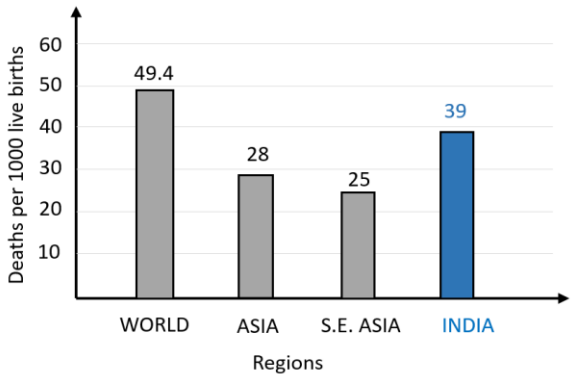

Fig 1: The Infant Mortality of World, Asia, South East Asia(S.E. Asia) and India respectively [5],[6],[7],[8]

There are 35,416 government hospitals in India. According to National Health Profile-2018, only 2% of the doctors serve in rural areas (which accounts for 68% of the population) [10]. The rest of the doctors (98%), however, are unwilling to move to rural regions that suffer from an acute shortage of medical practitioners [11]. The scarcity of healthcare centres in villages, a deficit of qualified doctors in India has been highlighted by WHO, where there are merely three qualified allopathic doctors and three nurses per every 10,000 population [12]. The WHO norm states that for every 1000 persons, one doctor should be available [13]. Fig 2 shows a typical diagnosis structure for a patient using healthcare facilities.

Additionally, MRI scans and imaging facilities are expensive and obtaining an accurate diagnosis is cumbersome [14]. Patients in fast-growing metropolitan cities have better access to diagnostic imaging facilities, whereas those in rural areas and low growth economic regions do not have easy access to diagnostic imaging facilities [15].

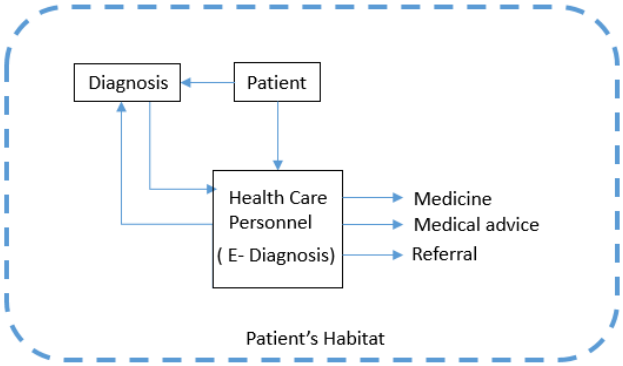

Fig 2: Diagnosis Structure

The rural population have low purchasing power, less number of quality medical infrastructure facilities, low private sector investment, and inadequacy of qualified physicians [15]. Despite these shortcomings, rural India has seen tremendous growth in the usage of smartphones and internet data connections (via smartphones). India is globally ranked second after China, in smartphone usage [16]. Smartphone usage in India is expected to increase from 26% in 2016 to 38% in 2021 [17]. There will be an increase in the percentage of internet users as well [18]. Moreover, India has the most affordable mobile data in the world. One GB of data costs a mere ₹18.5 (USD 0.26) as compared to the global average of ₹600 (USD 8.53) for the same [16]. This aids accessibility to the internet via smartphones, providing better opportunities for e-diagnosis.

This paper aims to develop a mobile-based application that detects pneumonia in its preliminary stage. This application aids in the diagnosis of pneumonia with the help of an X-Ray image of the patient. The app also can procure medical advice via e-diagnosis (over the internet). Further, a module for recording the patient's breathing sounds using a peripheral attachment to the smartphone, is proposed.

II. METHODOLOGY

A. Design Methodology

Data collection is done using 157 X-ray images, of which 73 are X-ray images of pneumonia patients while 84 are X-ray images of healthy individuals, through google search engine database, using a command-line Python program [19]. Subsequently, a refined dataset consisting of 84,495 X-ray images based on two image categories of pneumonia-infected and normal X-rays, is procured [20]. The collected data is filtered using a reference dataset from:- a) www.radiologyassistant.nl which is a trusted repository for radiographic imagery for pneumonia infections [21]; and b) The Chest X-Ray: A Survival Guide authored by Gerald de Lacey, Simon Morley and Laurence Berman [22]. The filtered data is used to train the Artificial Neural Network (ANN) by using MobileNets: Open-source models for on-device computer-vision based classification application for TensorFlow [23], an open-source machine learning (ML) platform for running ML models [24]. The result of training the ANN model is the generation of two files; namely a) output_graph.pb (model file), and b) output_labels.txt (label file, where the label is the name given to the images of normal and pneumonia patients). The Trained ANN is used to predict the confidence level of the labels with the help of a known X-ray image of a pneumonia patient to verify the accuracy of the model. On verification, the level of confidence would predict the stage and the possible complications of pneumonia in the patient (fig.4). A similar methodological framework could be followed to train the ANN for the breathing sound patterns of pneumonia affected persons. This ANN model is still under development and will be tested in the future.

Additionally, a module is developed to record the breathing patterns of pneumonia patients. This breathing pattern recorder module comprises a microphone, a diaphragm and an auxiliary audio input jack. This module has two variants (fig.3):- One is a low-cost wired module, and the other is a Bluetooth-based wireless module.

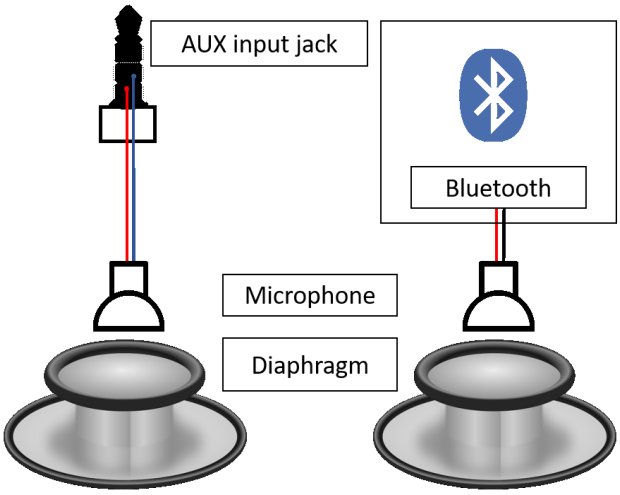

Fig 3: Wired module and Wireless module

B. Development of the mobile application

A mobile application is developed in Android Studio 3.3.1 using Java at the backend and XML for the frontend. A test app, obtained from TensorFlow library is modified by adding the label and model files. Furthermore, three functionalities, namely; patient profile, X-ray analyser, and Emergency Reporting, are added. The input to the application is an uploaded X-ray image which gives the confidence level of pneumonia prevalent in the patients. Additionally, there could be an option for uploading an audio file for analysing the breathing patterns of patients. An emergency reporting facility for pneumonia patients via e-diagnosis over the internet connection is also provided.

C. Development of Prototype

MobileNets for TensorFlow, being an on-device standalone analysis software (without needing to use the cloud computing platform), can work independently of an internet connection, which is an added advantage in the rural context. The X-ray image data is analysed with the help of the trained model in the ANN to predict the probability of pneumonia in the test data presented by the user. Also, preliminary advice on dealing with the disease is provided to the user. In the scenario where there is an active internet connection, the user can get connected to a group of expert medical practitioners (situated at distant metropolitan regions or anywhere in the world) available online, who can provide a real-time opinion for accurate prediction of the disease and affirm or disagree with the results provided by the

application. The system is developed in such a manner that there is always a pool of doctors online, to serve the users of this application.

To summarise, the confidence level of pneumonia is based on two labels defined as—'no pneumonia' or 'presence of pneumonia'—with the *stage* and *type* prediction. In addition, the results can be reaffirmed in coordination with medical experts available online who can act as an optional service to the users.

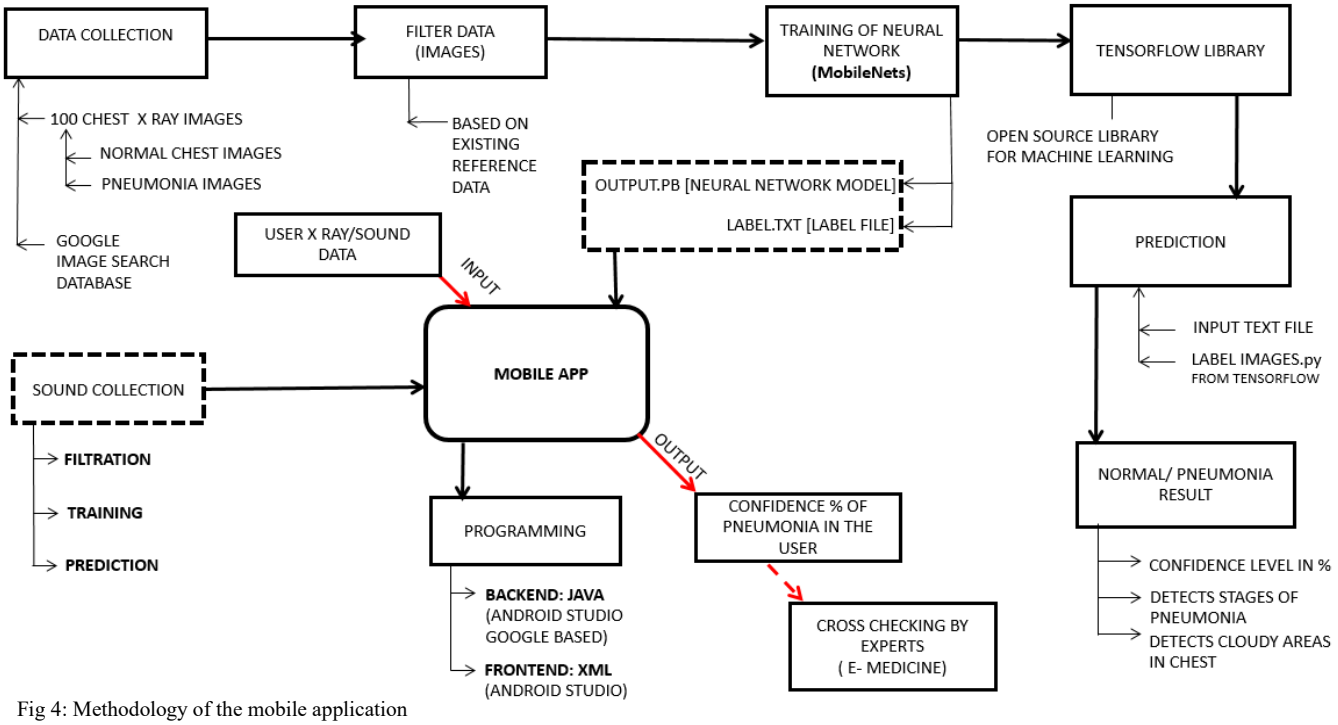

Fig 4: Methodology of the mobile application

The design methodology used in the development of this application for the preliminary detection of pneumonia is based on Artificial Neural Networks (ANN). ANN algorithm like MobileNets is faster in detection, with higher accuracy than Google inception v3. Complex algorithms are otherwise necessary to compare a single image with a large number of reference images. This method of comparison would require a large amount of computing power and time. Selection of MobileNets as an on-device classifier based on ANN is most favourable, as it is an open-source software from TensorFlow that works with minimal hardware and software requirements, which is generally met with smartphones available in the market today.

III. RESULTS AND DISCUSSION

A. *System Description*

The user needs to take a picture of the chest X-ray Image using his smartphone camera. Once this data is uploaded to the app, within a short period, the app can provide the results of the confidence percentage in the detection of the disease. The detection framework is implemented with the help of three layers—primary, secondary and tertiary. The primary layer is a filter that detects whether the patient has pneumonia or not. If pneumonia is detected, the secondary layer predicts the cause—viral or bacterial. The tertiary layer predicts the confidence level of pneumonia based on a fixed range. This per cent range 1–5, 20–30, and 80–100 denotes mild, intermediate, and severe pneumonia, respectively. The percentage of specificity shows the pneumonia type and confidence level of the prevalence of pneumonia.

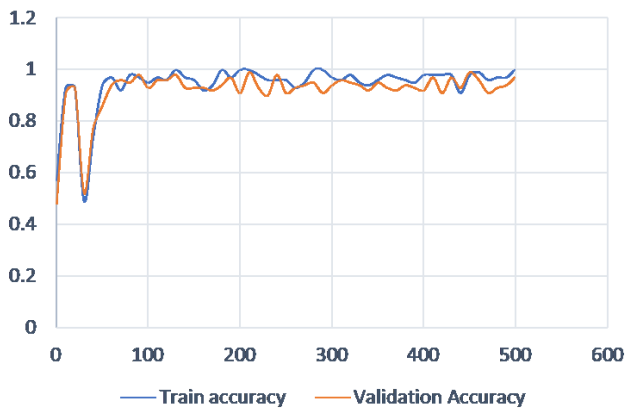

Fig 5: Accuracy of the ANN model (X-ray) in training and validation

This problem is addressed with the help of the solution presented in this paper, where a patient suspected of pneumonia or other respiratory illnesses is advised to take an X-ray. For this application to function correctly, it is a prerequisite that the patient captures the image of the X-ray film against a bright white background.

B. *Accuracy and Cross entropy*

The accuracy of training the ANN model based on X-ray image analysis is defined as the number of images precisely classified to the whole batch of images. Validation accuracy is the precision of correctly classified images calculated on a random selection of images from a different batch, from the

previously used dataset. The total accuracy of this MobileNets model is found to be 96.8%, which is statistically significant (fig 5).

Cross entropy gives an analysis of how good the model is learning and is a loss function. Fig 6 shows the cross-entropy of the model in training.

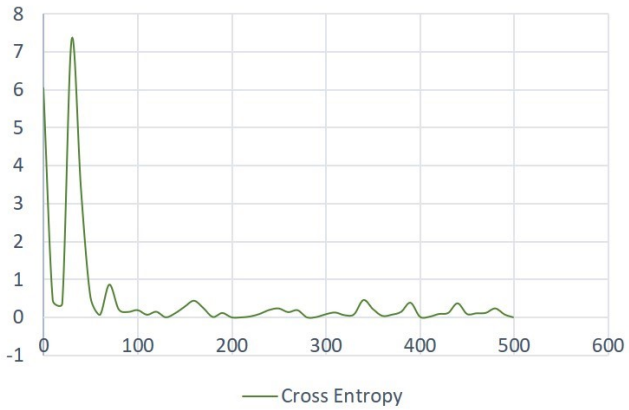

Fig 6: Cross entropy of the ANN (X-ray) model in training

The retrain.py is a python script that has been sourced from the TensorFlow library. The script is used for training the model, by defining a training step size of 500, MobileNets architecture v1, and defining the output directory for storing the output_graph.pb and output_labels.txt, as shown in Table I. Table II shows the CMD script for validating the model using the known pneumonia image dataset to cross-check the accuracy of the model.

TABLE I. CMD SCRIPT FOR TRAINING THE MODEL

```
python retrain.py
--image dir ./X-Rays
--bottleneck dir=./output/bottlenecks --
how_many_training_steps=500
--model_dir=./output/models/
--summaries_dir=
          ./output/training summaries
          /mobilenet 1.0 224
--architecture= mobilenet 1.0 224
--output_graph ./output/output_graph.pb
--output_labels ./output
          /output_labels.txt
--saved_model_dir ./output/saved_mode
```

TABLE II. CMD SCRIPT FOR VALIDATING THE MODEL

```
python label image.py
--graph=./output/output_graph.pb
--labels=./output/output_labels.txt
--image=./test X-Ray.jpg
```

C. Discussions

The ability of the app to detect the early stages of pneumonia is convenient and affordable to users in the long run. The design's integrated X-ray image analysis employing machine learning, TensorFlow, and deep learning helps in quick diagnosis and enables the users to cross-check information from medical sources. Fig 7 shows the various screens of the smartphone application.

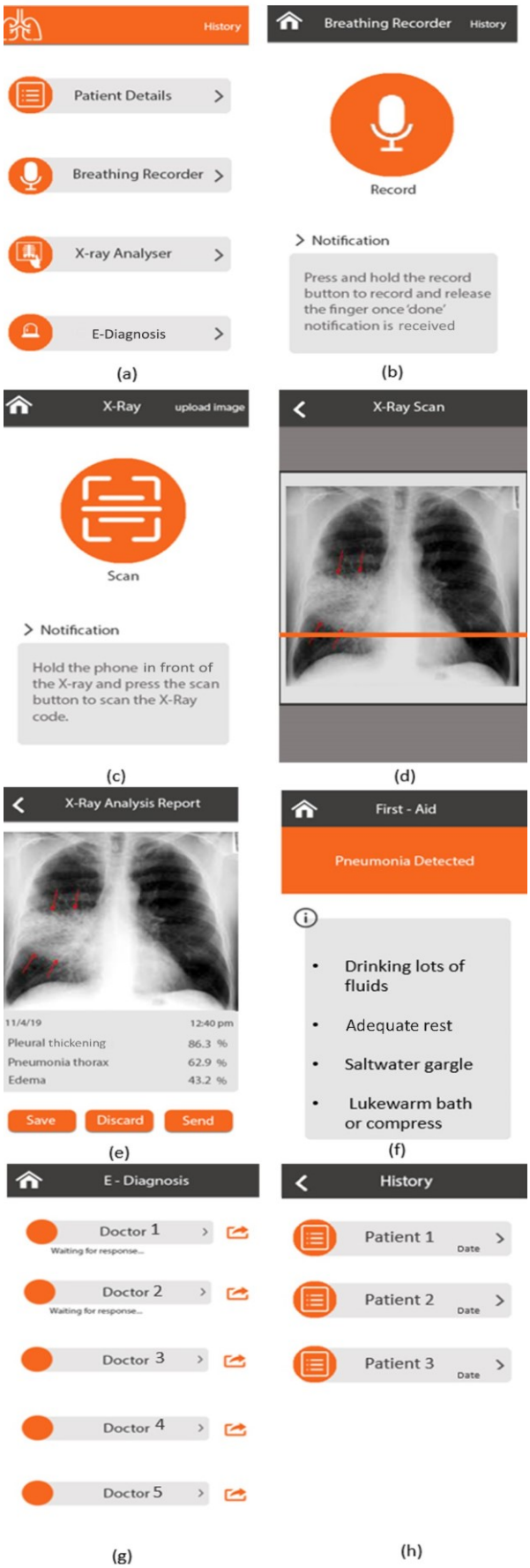

Fig.7: Screens of the mobile application

A peripheral module to record breathing data is developed (fig 3). The breathing pattern, while inhaling and exhaling are recorded through the module connected to a smartphone via wired or wireless means. Consequently, a model similar to the ANN model developed for X-ray image analysis could be developed. Integration of this model to the existing app would increase the precision of pneumonia prediction based on an increase in the number of input parameters. The input parameters in the proposed system would be a) X-ray image (developed in this study), and b) breathing pattern (yet to be developed).

IV. Conclusion

Pneumonia is a critical respiratory infection which could become life-threatening if medical aid is not administered in optimum time. Diagnostic facilities are expensive and not affordable to the rural communities in the developing world. Moreover, there is an acute shortage of hospitals and doctors in rural regions, which makes the early detection and diagnosis, a difficult task for most people in rural areas.

After an X-ray image of the chest is obtained by a patient, there is a possibility of long waiting periods, before meeting concerned doctors. This delay may aggravate the health conditions of pneumonia patients. This smartphone application can be of significant help to any individual, including infants and children. This app requires a chest X-ray image and a smartphone for early detection of pneumonia. It is also convenient and affordable to users, as it helps save their time and money. The X-ray image analysis helps in improving the prediction accuracy and provides confidence to the patients, to approach concerned medical facilities for necessary treatment at the earliest. The patients have an additional option to use the e-diagnosis facility for cross-examining the authenticity of the app's prediction and also avail guidance and support from professional healthcare experts available online. In future, breathing pattern analysis would be integrated into the app (employing an ANN model) which could increase the accuracy of predictions. Furthermore, validation, deployment and real-case testing may be undertaken. This application can help in the battle against pneumonia, and induce awareness about the early detection and treatment of pneumonia in rural regions.

References

[1] American Lung Association Scientific and Medical Editorial Review Panel, "Lung Health & Diseases," American Lung Association, 2019. [Online]. Available: https://www.lung.org/lung-health-and-diseases/lung-disease-lookup/pneumonia/learn-about-pneumonia.html. [Accessed: 18-Apr-2019].

[2] NHS, "What causes pneumonia?," NHS, 2016. [Online]. Available: https://www.nhs.uk/conditions/pneumonia/. [Accessed: 14-May-2019].

[3] Peter Crosta and M. Judith Marcin, "What you should know about pneumonia," MedicalNewsToday, 2017. [Online]. Available: https://www.medicalnewstoday.com/articles/151632.php. [Accessed: 18-Apr-2019].

[4] unicef, "PNEUMONIA," unicef, 2018. [Online]. Available: https://data.unicef.org/topic/child-health/pneumonia/. [Accessed: 15-Apr-2019].

[5] index mundi, "India Infant mortality rate," CIA World Factbook, 2018. [Online]. Available: https://www.indexmundi.com/india/infant_mortality_rate.html. [Accessed: 16-Apr-2019].

[6] IGME, "UN Inter-agency Group for Child Mortality Estimation," IGME, 2018. [Online]. Available: https://childmortality.org/data. [Accessed: 15-Apr-2019].

[7] Statista, "Infant mortality rates in South East Asian countries between 2010 and 2015 (infant deaths per 1,000 live births)," Statista, 2019. [Online]. Available: https://www.statista.com/statistics/590013/infant-mortality-rates-in-south-east-asia/. [Accessed: 16-Apr-2019].

[8] United Nations Population Division | Department of Economic and Social Affairs, "World Mortality 2017: Data Booklet (ST/ESA/SER.A/412)," 2017.

[9] WHO, "World Pneumonia Day 2018," WHO, 2019. [Online]. Available: https://www.who.int/maternal_child_adolescent/child/world-pneumonia-day-2018/en/. [Accessed: 16-Apr-2019].

[10] Ministry of Health and Family Welfare, "National Health Profile 2018 - Birth Rate," p. 23, 2018.

[11] R. Nagarajan, "6 states have more doctors than WHO's 1:1,000 guideline," TNN, 2018. [Online]. Available: https://timesofindia.indiatimes.com/india/6-states-have-more-doctors-than-whos-11000-guideline/articleshow/65640694.cms. [Accessed: 14-Apr-2019].

[12] K. D. Rao, R. Shahrawat, and A. Bhatnagar, "Composition and distribution of the health workforce in India: estimates based on data from the National Sample Survey," WHO South-East Asia J. public Heal., vol. 5, no. 2, p. 133, 2016.

[13] Global Health Observatory, "Density of physicians (total number per 1000 population, latest available year)," WHO, 2019. [Online]. Available: https://www.who.int/gho/health_workforce/physicians_density/en/. [Accessed: 14-May-2019].

[14] Medifee Healthcare, "MRI Scan – Importance, Procedure & Cost in India," Medifee Healthcare, 2019. [Online]. Available: https://www.medifee.com/blog/mri-scan-importance-procedure-and-cost-in-india/. [Accessed: 18-Apr-2019].

[15] Bhaskar Vittal, "Access All Areas: Diagnostic Imaging in Emerging Economies," GlobalData, 2010. [Online]. Available: https://www.hospitalmanagement.net/features/feature79325/. [Accessed: 15-Apr-2019].

[16] Economic Times, "India has the cheapest mobile data in world: Study," Economic Times, New Delhi, Mar-2019.

[17] I. Traffic, "India - 2021 Forecast Highlights," pp. 1–5, 2016.

[18] T. M. Tm, "21st edition ICUBE Digital adoption & usage trends ICUBE TM : Internet In India Asia ' s largest authoritative resource for Digital Marketing and Planning Decisions," pp. 1–10.

[19] GitHub, "Google Images Download," GitHub, Inc., 2019. [Online]. Available: https://github.com/hardikvasa/google-images-download. [Accessed: 13-May-2019].

[20] M. G. Daniel Kermany, Kang Zhang, "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification," Mendeley, 2018. [Online]. Available: https://data.mendeley.com/datasets/rscbjbr9sj/2. [Accessed: 14-May-2019].

[21] H. A. Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto and (Submitted on 17 Apr 2017), "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," Cornell University, 2017. [Online]. Available: https://arxiv.org/abs/1704.04861. [Accessed: 10-May-2019].

[22] M. A. Gerald de Lacey, S. Morley, and L. Berman, The chest X-ray: a survival guide. Elsevier Health Sciences, 2008.

[23] Andrew G. Howard and M. Zhu, "MobileNets: Open-Source Models for Efficient On-Device Vision," Google AI blog, 2017. [Online]. Available: https://ai.googleblog.com/2017/06/mobilenets-open-source-models-for.html. [Accessed: 14-Apr-2019].

[24] Tensorflow, "An end-to-end open source machine learning platform," GitHub, 2019. [Online]. Available: https://www.tensorflow.org/. [Accessed: 14-Apr-2019].